

# Scalable HEVC 표준 기술 동향

김경혜, 조현호, 심동규  
광운대학교

## 요약

네트워크 기술과 멀티미디어 압축 기술이 발달함에 따라 사용자들은 고화질의 멀티미디어 콘텐츠를 무선 환경에서도 끊김 없이 제공 받기를 원하고 있다. 또한, 스마트폰과 스마트패드와 같은 다양한 멀티미디어 장치가 일상생활에서 보편적으로 사용되면서 디바이스 간에 끊김 없는 미디어의 연동을 가능케 하는 N-스크린 서비스가 점차 확산되고 있다. 이러한 시장의 흐름에 발맞춰 JCT-VC (Joint Collaborative Team on Video Coding)는 차세대 동영상 압축 표준 기술인 HEVC (High Efficiency Video Coding)를 바탕으로 공간적 (Spatial), 화질적 (Temporal) 스케일러빌리티 (Scalability)를 제공하는 확장 표준인 SHVC (Scalable HEVC)를 표준화 중에 있다. JCT-VC는 SHVC 부호화 기술들을 평가하기 위하여 SCE (Scalable Core Experiment)를 진행하고 있으며, SHVC 표준은 2014년 6월에 표준화가 완료 될 것으로 예상된다. 본 고에서는 비디오 스케일러빌리티의 기본 개념 및 SHVC 표준화 과정에서 제안된 주요 부호화 기술들과 이슈들을 소개한다.

## I. 서론

최근 스마트기기의 보급과 네트워크 관련 기술의 발전에 따라 실시간으로 동영상을 스트리밍 받는 장치들이 늘어나고 있다. 지금까지 콘텐츠 제공자들은 다양한 환경을 고려한 서비스를 위하여 각 환경에 최적화된 비트스트림을 생성하고 이를 제공하는 방식을 사용하였다. 이러한 방식을 simulcast 라고 하는데, 이 simulcast 방식은 동일 영상에 대해서 서로 다른 부호화 조건을 사용하여 다수의 비트스트림을 생성하기 때문에, 콘텐츠 서버의 용량을 증가시키는 단점을 가진다. 이러한 방법은 각 단말기의 환경을 고려한 최적의 비트스트림을 생성할 수 있으며, 시스템의 구축이 상대적으로 간단하여서 지금까지 널리 사용되고 있다. 하지만, 최근 본격적으로 4K-UHD (3840

×2160)와 8K-UHD (7680×4320) 급의 영상에 대한 취득 및 재생 장치들이 보급되기 시작했으며, VGA (640×480)급부터 Full-HD (1920×1080)급 해상도를 갖는 다양한 스마트기기들이 이미 널리 사용되고 있기 때문에 기존의 simulcast 방식으로는 다수의 단말기에 대해서 효과적인 콘텐츠의 제공이 점점 더 어려워질 것으로 예상된다. 또한, 다양한 멀티미디어 재생 장치의 보급은 N-스크린 서비스를 확산시킬 것으로 예상하는데 기존의 simulcast 방식으로는 N-스크린 서비스를 효과적으로 구현할 수 없다는 한계가 있다.

콘텐츠를 이용하는 사용자들의 단말기 성능이나 네트워크 상황, 또는 단말기의 해상도 등에 실시간으로 대응할 수 있는 영상 압축 방법으로 스케일러블 영상 코딩 (Scalable Video Coding, SVC)이 있다. SVC는 시간적 (Temporal), 공간적 (Spatial), 그리고 화질적 (Quality or SNR) 계위성을 갖도록 영상을 단일의 비트스트림으로 압축하는 방식이다. SVC를 이용하면 <그림 1>과 같이 단말기 디스플레이 패널의 크기, 네트워크 대역폭 (Bandwidth) 등 가변적으로 변하는 사용자의 환경을 고려하여 이미 부호화 된 단일 비트스트림으로부터 일부의 비트스트림만을 추출하여 해당 환경에 대해 효과적으로 대응할 수 있다. SVC 기술을 사용하여 영상을 압축하면 기존의 simulcast 방식이 제공하는 기능을 그대로 제공할 수 있을 뿐만 아니라 simulcast 방식에 비해서 압축률도 높아서 콘텐츠 서버의 용량 문제도 해결할 수 있다. 이러한 SVC 기술은 이미 MPEG-2와 H.264/AVC 등의 표준에 대한 확장 표준[1]으로 제정되었으나 앞서 언급한 simulcast 방식의 장점으로 인하여 실제로 서비스에 응용된 사례는 많지 않았다.

2013년 1월 HEVC (High Efficiency Video Coding)이라는 차세대 동영상 압축 표준 기술이 제정되면서 HEVC를 표준화하였던 JCT-VC (Joint Collaborative Team on Video Coding)는 HEVC를 기반으로 SVC 표준 기술을 개발하고 있다. HEVC 기반의 SVC 표준 기술인 Scalable HEVC (SHVC)는 최근의 멀티미디어 환경을 고려할 때, 기존의 MPEG-2, H.264/AVC 기반의 SVC에 비해서 널리 사용될 수 있을 것으로 예상된다. 본 기고에서는 스케일러블 영상 압축에 대한 개념을 살펴보고,

HEVC 기반의 스케일러블 영상 압축 표준 기술인 SHVC의 표준 기술 및 표준화 동향을 살펴본다.

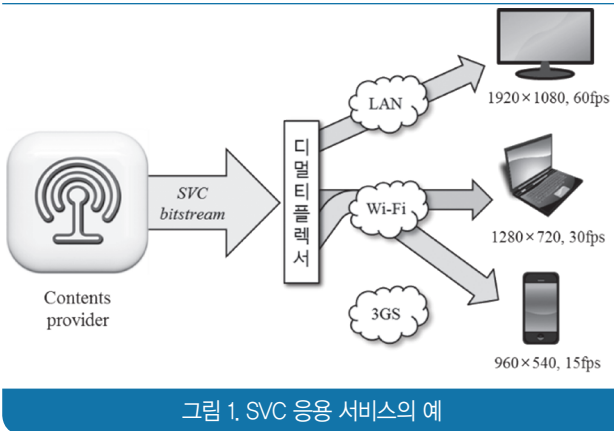


그림 1. SVC 응용 서비스의 예

## II. SVC의 스케일러빌리티 (Scalability)

앞에서 설명한 바와 같이 SVC는 공간적, 시간적, 그리고 화질적 측면에서 스케일러빌리티를 제공한다. SVC의 공간적 스케일러빌리티를 이용하면 부호화 된 단일의 비트스트림만으로도 각 단말기 장치의 환경에 맞춰 해상도를 조절할 수 있다. 시간적 스케일러빌리티를 이용하면 각 단말기 장치의 환경 또는 네트워크 상태에 적응적으로 fps (frame per second)를 조절할 수 있으며, 화질적 스케일러빌리티를 사용하면 화질을 조정하여 서비스하는 것이 가능하다. 본 장에서는 SVC에서 지원하는 3가지의 스케일러빌리티에 대해 자세히 알아본다.

### 1. 스케일러빌리티 (Scalability)

〈그림 2〉는 SVC의 시간적 (Temporal), 공간적 (Spatial), 그리고 화질 (Quality) 계층성을 나타낸 것이다. 〈그림 2〉에서 각 블록은 SVC를 통해 부호화되는 단일 비트스트림의 일부 데이터를 의미한다. SVC는 simulcast와 달리 영상을 압축한 후 스케일러빌리티를 지원하는 단일의 비트스트림을 출력하는데, 해당 비트스트림은 〈그림 2〉와 같이 지원하는 스케일러빌리티에 따라서 세부적으로 분할이 가능한 구조를 가진다. 예를 들어, 〈그림 2〉에서 A의 비트스트림만을 전송하면 데이터를 전송 받은 단말장치에서는 SD급 해상도, 7.5fps, 일반 화질의 비디오를 복호화하여 재생할 수 있다. 이때 추가적으로 B에 해당하는 데이터를 더 전송 받은 경우에는 HD급 해상도의 비디오를 복호화할 수 있게 된다. 마찬가지로 C, D를 더 전송 받게 되는 경우 HD 급의 해상도 영상에 대해서 30fps까지 프레임 율을 증가시

킬 수 있으며, E, F를 더 전송 받게 되면 HD 해상도, 30fps에서 고화질의 비디오를 복호화하고 이를 재생할 수 있는 것이다.

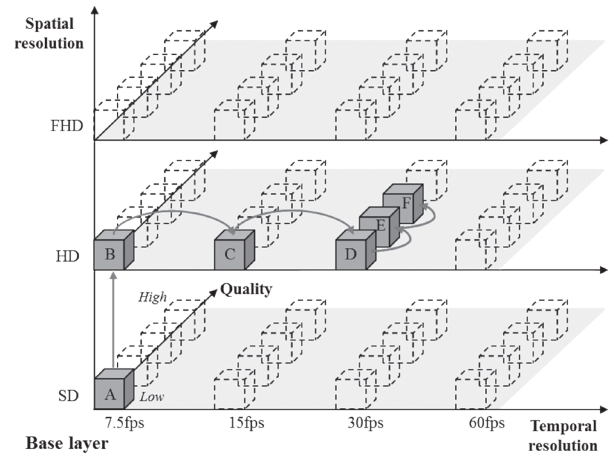


그림 2. SVC의 스케일러블 구조에 대한 개념도

### 1.1 시간적 (Temporal) 스케일러빌리티

SVC는 단일의 비트스트림으로 재생하려는 영상의 fps를 조절할 수 있는 시간적 스케일러빌리티를 지원하며, 이를 위해서 〈그림 3〉과 같은 hierarchical-B 참조 구조를 사용한다. 〈그림 3〉은 HEVC에서 사용하는 hierarchical-B 참조 구조를 간단히 나타낸 것이다. 이러한 구조에서는 임의의 프레임은 기준으로 하위에 위치한 프레임은 참조할 수 있으며, 상위에 위치한 프레임들은 참조하지 않는다. 예를 들어, B<sub>1</sub> 프레임들은 하위에 위치한 B<sub>0</sub>, I, P 프레임들을 참조할 수 있지만, B<sub>2</sub> 프레임은 참조하지 않는다. 디코더 관점에서 다시 살펴보면, B<sub>1</sub> 프레임을 복호화 하는 데 B<sub>0</sub> 프레임이 필요하지 않기 때문에 이에 해당하는 비트스트림을 버려도 복호화 하는 데 문제가 생기지 않는다. 이러한 계층적 참조 구조를 이용하면 영상의 프레임율을 다양하게 변경할 수 있다. 예를 들어, 현재 영상의 프레임율이 60fps 인 경우, 단말기에서 B<sub>2</sub> 프레임들을 제외한 나머지 비트스트림만 추출하여 복호화하면 30fps의 영상을 얻을 수 있다. 또한 I, P, B<sub>0</sub> 프레임만 복호화하여 사용하면 15fps의 영상 재생이 가능하다.

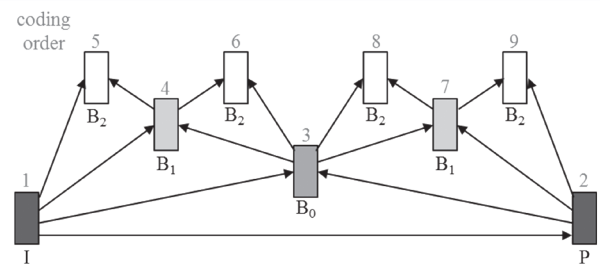


그림 3. HEVC의 hierarchical-B 참조 구조

## 1.2 공간적 (Spatial) 스케일러빌리티

SVC는 하나의 비트스트림으로 다양한 해상도의 영상을 제공할 수 있는 공간적 스케일러빌리티를 지원한다. <그림 4>는 공간적 스케일러빌리티를 이용한 응용 서비스의 예이다. 다양한 해상도를 갖는 영상들은 SVC 부호화기를 통해 하나의 비트스트림으로 만들어진다. 사용자 단말기에서는 단말기가 지원하는 해상도에 맞게 필요한 비트스트림을 추출하여 복호화 하고 재생할 수 있다. 예를 들어, <그림 4>와 같이 세 가지 해상도의 영상을 SVC를 이용해 부호화하여 전송할 때, 노트북에서는 기본 계층과 향상 계층 1에 해당하는 비트스트림만 있으면 노트북에서 지원하는 해상도 (1280×720)의 영상을 재생할 수 있다. 스마트폰을 사용 중인 경우에는 기본 계층의 비트스트림만 추출하여 복호화하면 스마트폰 해상도에 맞는 영상을 재생할 수 있다. SVC는 해상도가 서로 다른 각 계층의 비디오를 코딩 할 때 서로 독립적으로 부호화하는 simulcast와 달리, 부호화하려는 계층의 하위 계층 정보를 참조하여 부호화하기 때문에 압축 효율이 높다는 장점을 갖는다.

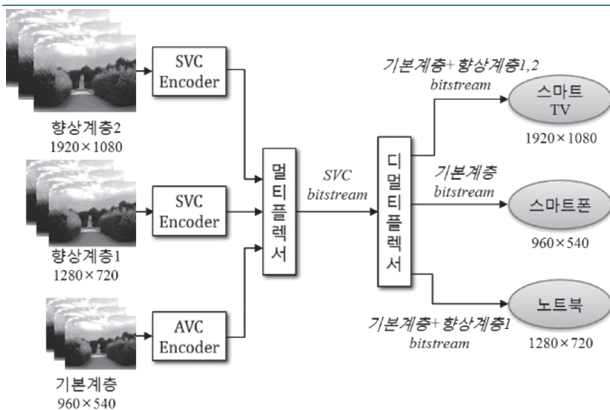


그림 4. 공간적 스케일러빌리티 응용 서비스의 예

## 1.3 화질적 (SNR or Quality) 스케일러빌리티

SVC는 단일의 비트스트림으로부터 서로 다른 화질을 갖는 영상을 제공할 수 있는 화질적 스케일러빌리티를 지원한다. 이러한 화질적 스케일러빌리티를 사용하면 부호화된 단일 비트스트림을 이용하여 네트워크 상황에 맞춰 제공하는 콘텐츠의 화질을 조정하는 것이 가능하다. 예를 들어 야외에서 스마트폰을 이용해 고화질의 스포츠 중계 서비스를 이용하다가 지하로 이동할 경우 네트워크 수신률이 낮아질 수 있다. 이 때, 고화질의 영상을 계속 전송하게 될 경우, 즉 많은 양의 데이터를 전송할 경우 단말기에서는 데이터 수신이 원활하지 않아 영상을 제대로 재생할 수 없다. 하지만 화질적 스케일러블 부호화를 이용하면 네트워크 상황이 변함에 따라 영상이 끊기는 문제를 해결할 수 있다. 고화질의 영상을 전송하다가 데이터 전송 및 수신이 원활

하지 않을 경우 적은 용량, 낮은 화질에 대한 비트스트림만 전송하면 단말기에서는 화질은 떨어지지만 끊기지 않는 영상을 받아볼 수 있다.

<그림 5>는 두 가지 화질을 제공하는 SVC의 화질적 스케일러빌리티의 부호화 및 응용 서비스의 예를 보여준다. 화질적 스케일러빌리티는 동일한 해상도의 영상을 다른 화질로 부호화 하는 것이므로 모든 계층의 입력 영상 해상도는 동일하다. 부호화 하는 영상의 화질은 QP (Quantization Parameter) 값을 이용해 조절할 수 있으며, QP 값이 높을수록 낮은 화질로 부호화된다. 부호화 된 비트스트림은 서비스를 이용하는 네트워크의 대역폭에 따라 선택적으로 전송할 수 있다. Wi-Fi를 사용하는 경우에는 고화질 영상에 대한 비트스트림을 전송하고, 상대적으로 낮은 대역폭을 갖는 3GS를 사용 중인 경우 기본 화질 (저화질)의 비트스트림만 보내줄 수 있다.

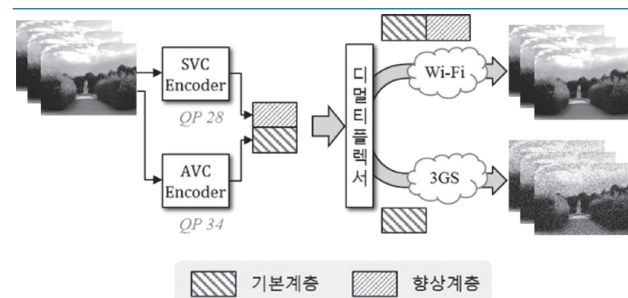


그림 5. 화질적 스케일러빌리티 부호화 및 응용 서비스의 예

## III. SHVC 표준화 동향

HEVC 표준이 완료되어감에 따라, 2012년 7월 JCT-VC 10 차 회의에서 HEVC 기반 스케일러블 부호화 확장 표준을 위한 Final Call for Proposal[2]이 발행되었다. 2013년 1월 JCT-VC 11차 회의에서는 HEVC Final Draft International Standard (FDIS)를 발행하였으며, HEVC 확장 표준인 SHVC의 test model 경쟁이 이루어졌다. 삼성, Vidyo, Nokia 등 여러 기관에서 SHVC test model 경쟁에 참여하였으며, 표준화 회의에서는 각각의 모델에 대한 전체적인 성능 비교를 하였다. 하지만 모델 별로 사용한 부호화 기술들이 매우 다양해서 정확한 성능 평가를 하기엔 어려움이 있었기 때문에, 기본적인 계층 간 텍스트처 예측 기술만 포함하여 SHVC 소프트웨어 초기 모델을 결정하였고, test model로 제안되었던 스케일러블 코덱의 부호화 기술 및 성능 평가를 위해 툴 테스트 (Tool Experiments, TE)를 진행하였다. 이 중 주요 기술들을 바탕으로 현재 SHVC 표준화 회의에서는 SCE (Scalable Core Experiments)를 진행



중이며, SCE 관련 기술에 대해서는 5장에서 자세히 살펴볼 예정이다.

SHVC 소프트웨어 초기 모델은 TextureRL과 RefIdx 방식을 모두 지원하였다. TextureRL은 현재 코딩중인 향상 계층의 블록에 대해서 계층 간 예측 시 필요한 참조 계층의 블록을 가져오는 방식이고, RefIdx는 어느 한 시점에서 향상 계층 픽처를 코딩 할 때, 시간적으로 동일한 위치에 있는 참조 계층의 픽처를 향상 계층의 참조 리스트에 저장하여 계층 간 예측을 수행할 수 있도록 하는 방법이다[3][4]. SHVC 표준화 회의에서는 두 가지 방식 중 하나를 선택하기 위한 회의가 진행되었으며, JCT-VC 13 차 회의에서 RefIdx를 사용하는 것으로 결정되었다[5]. 현재 SHVC 참조 소프트웨어 (SHM)는 2.1 버전까지 릴리즈 되었으며[6], 이는 HEVC 소프트웨어인 HM-10.1을 기반으로 하고 있다.

SHVC는 여러 계층의 영상을 코딩 할 뿐만 아니라 향상 계층의 부호화 성능을 증가시키기 위한 다양한 툴들이 추가되기 때문에 HEVC 대비 부복호화기의 복잡도가 매우 높다. 따라서 SHVC 표준화 회의에서는 기고되는 기술들에 대해 복잡도, 연산량, 메모리 사용량 등을 세세하게 검토하여 기술의 채택 여부를 결정하고 있다. 이에 따라서 표준화 회의에서는 SHVC의 SCE 관련 기술 및 부호화 성능 향상을 위한 기술과 더불어 툴의 복잡도를 줄이는 방법에 대한 기술들이 기고되고 있으며, 2014년 6월에 최종적으로 SHVC 표준을 완료할 것으로 예상된다.

#### IV. SHVC 부복호화기 구조

〈그림 6〉은 2개의 계층을 갖는 SHVC 부호화기의 구조를 나타낸 것이다. 계층은 크게 기본 계층 (Base layer)과 향상 계층 (Enhancement layers)로 나눌 수 있으며, 향상 계층은 여러 개가 존재할 수 있다. 최상위 계층을 제외한 하위 계층들 (기본 계층 포함)은 모두 최상위 계층으로부터 공간적 또는 화질적으로 다운 스케일 된 영상들이다. 기본 계층은 단일 시점 디스플레이와의 호환성을 위하여 추가적인 툴 없이 HEVC와 동일한 방법으로 부호화 되는 계층을 의미한다. 향상 계층 부호화기는 HEVC를 기반으로 하고 있으며, 하위 계층 정보를 이용하여 부호화하는 기술인 계층 간 예측 툴 (Inter-layer prediction tool)을 포함하고 있다. 이때, 향상 계층의 부호화를 위해 참조 되는 하위 계층을 참조 계층이라 부르기도 한다. 향상 계층은 계층 간 예측 툴을 사용하기 때문에 하위 계층과의 의존성 (dependency)이 존재한다. 따라서 향상 계층을 부복호화 하기 전에, 먼저 해당 계층의 하위 계층들을 모두 부복호화 해야 한

다. 계층 간 예측 툴에는 크게 하위 계층의 텍스처를 참조하는 기술과 움직임 정보를 참조하는 기술이 있으며, 이에 대해서는 다음 장에서 상세히 다루도록 하겠다.

RefIdx 방식을 사용하고 있는 SHVC는 참조 계층의 영상을 향상 계층의 참조 리스트에 저장하여 향상 계층에서 참조 계층을 이용한 예측을 수행할 수 있도록 하고 있다. 향상 계층은 참조 리스트에 추가된 참조 계층의 영상으로부터 텍스처 및 움직임 정보 등을 참조하여 부호화 될 수 있다. 현재 SHVC에서 참조 계층은 향상 계층 참조 리스트의 맨 끝에 위치하게 되며, 이는 〈그림 7〉에 나타난 바와 같다. 현재 코딩 중인 향상 계층 프레임이 E2이고 E2가 참조하는 프레임은 E0, E4, E8이라고 가정하면, E0는 참조 리스트 L0에, E4와 E8은 참조 리스트 L1에 들어가게 된다. E2와 시간적으로 동일한 위치에 존재하는 참조 계층 프레임인 B2는 향상 계층의 프레임들로 이루어진 참조 리스트 구성이 끝난 후, 두 참조 리스트 L0, L1의 맨 뒤에 위치하게 된다. 공간적 스케일러빌리티의 경우, 향상 계층과 참조 계층의 해상도가 다르기 때문에 참조 계층 픽처를 향상 계층 해상도에 맞게 업 샘플링 한 후에 참조 리스트에 추가한다.

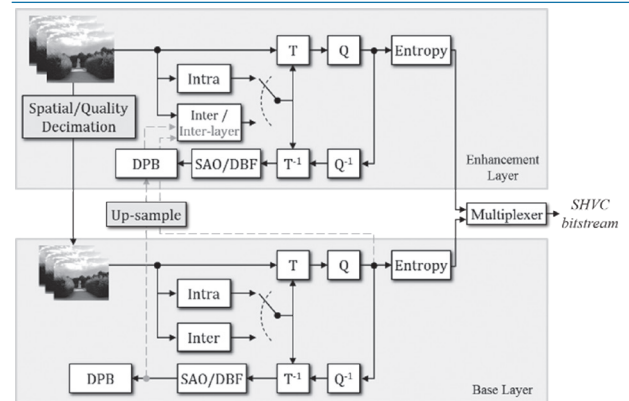


그림 6. SHVC 부호화기 구조도

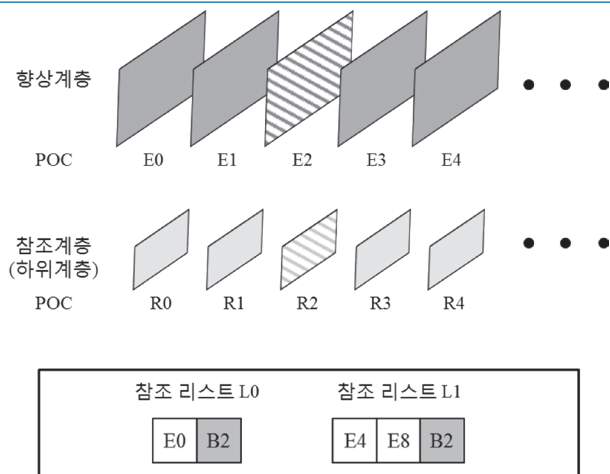


그림 1. 4D 아키텍처

본 장에서는 RefIdx 방식을 사용하고 있는 SHVC의 기본 구조에 대해 살펴보았다. 다음 장을 통해서 SHVC에서 사용되고 있는 부호화 기술 및 표준화 회의에서 이슈가 되고 있는 기술들에 대해 좀 더 자세히 알아보도록 한다.

## V. SHVC 부호화 기술

본 장에서는 현재 SHVC에서 사용하고 있는 계층 간 예측 기술과 SCE에서 다루어지고 있는 부호화 기술들에 대해 자세히 알아본다. 또한 다음 표준화 회의에서 논의될 SCE에 대해서도 간략히 살펴보도록 한다.

### 1. 계층 간 예측 (Inter-layer prediction)

스케일러블 영상의 각 계층은 해상도 또는 화질의 차이만 있을 뿐, 동일한 장면 (scene)의 영상이기 때문에 서로 간의 유사성이 매우 높다. 이러한 점을 이용하여 스케일러블 영상 부호화에서는 참조 계층으로부터 텍스처와 움직임 정보, 잔여신호 (residual) 등의 다양한 정보를 이용하여 향상 계층을 부호화하는 계층 간 예측 기술을 사용하고 있다. 현재 SHVC에서는 계층 간 예측 시 하위 계층의 텍스처와 움직임 정보를 참조하고 있으며, 이에 대해서 자세히 살펴보도록 하자.

#### 1.1 계층 간 텍스처 예측 (Inter-layer texture prediction)

스케일러블 영상 부호화에서 향상 계층 부호화 시, 시간적으로 동일한 시점에 존재하는 하위 계층의 텍스처를 참조할 수 있으며, 이러한 방법을 계층 간 텍스처 예측이라고 한다. 현재 SHVC에서는 코딩 중인 향상 계층 블록에 대해 동일한 위치에 있는 참조 계층의 블록을 예측 블록으로 사용하고 있다. 참조 계층의 픽처는 향상 계층의 참조 리스트에 저장된 후에 계층 간 텍스처 예측에 사용된다[7]. 공간적 스케일러빌리티의 경우에는 계층마다 해상도가 다르기 때문에, 참조 계층의 텍스처를 그대로 사용하지 않고 향상 계층의 해상도에 맞게 업 샘플링 한 후 사용한다. 예를 들어, 2배 공간적 스케일러빌리티의 경우 향상 계층과 참조 계층의 해상도는 가로, 세로 각각 2배씩 차이가 난다. 따라서 참조 계층의 텍스처를 가로, 세로 2배 업 샘플링 한 후에 향상 계층의 참조 리스트에 저장한다. 텍스처 업 샘플링은 DCT기반 보간 (interpolation) 필터를 사용하고 있으며, 휘도와 색차 성분에 대한 필터 계수는 각각 <표 1>, <표 2>와 같다[8]. 현재는 2배와 1.5배에 대한 필터 계수만 정의되어 있으며, 추후 표준화 회의에서 다양한 배율 (arbitrary ratio)에

대한 필터 계수를 결정할 것으로 보인다.

<그림 8>은 공간적 스케일러빌리티에서의 계층 간 텍스처 예측 방법을 간략히 나타낸 것이다. 향상 계층의 해상도가 1920×1080이고 참조 계층의 해상도가 960×540인 경우, 참조 계층은 우선 향상 계층의 해상도에 맞게 업 샘플링 된다. 향상 계층의 코딩 중인 현재 블록에 대하여, 업 샘플링 된 참조 계층 내의 동일한 위치에 존재하는 블록이 예측 블록으로 사용된다. 계층 간 텍스처 예측을 하는 데 있어서 반드시 동일한 위치의 참조 계층 블록을 예측 블록으로 사용하지 않고, 움직임 예측 (motion estimation)을 수행하여 참조 계층 내에서 최적의 블록을 찾을 수도 있으나, 현재 SHVC 소프트웨어에서는 동일한 위치의 블록만 사용하도록 구현되어 있다.

표 1. 휘도 성분에 대한 업 샘플링 필터 계수

Phase	Filter coefficients							
0	0	0	0	64	0	0	0	0
5/16 (1.5x)	-1	4	-11	52	26	-8	3	-1
8/16 (2x)	-1	4	-11	40	40	-11	4	-1
11/16 (1.5x)	-1	3	-8	26	52	-11	4	-1

표 2. 색차 성분에 대한 업 샘플링 필터 계수

Phase	Filter coefficients			
0	0	64	0	0
4/16 (1.5x)	-4	54	16	-2
5/16 (1.5x)	-6	52	20	-2
6/16 (2x)	-6	46	28	-4
8/16 (2x)	-4	36	36	-4
9/16 (1.5x)	-4	30	42	-4
11/16 (1.5x)	-2	20	26	-6
14/16 (2x)	-2	10	58	-2
15/16 (1.5x)	0	4	62	-2

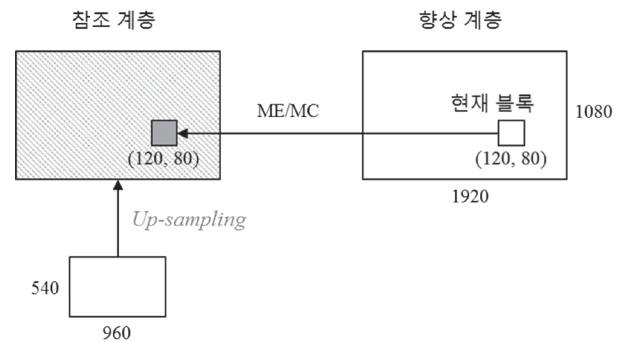


그림 8. 공간적 스케일러빌리티에서 계층 간 텍스처 예측의 예

### 1.2 계층 간 움직임 예측 (Inter-layer motion prediction)

계층 간 움직임 예측은 하위 계층과의 움직임 정보가 유사할 가능성이 높다는 점을 이용해 참조 계층의 움직임 정보를 사용하여 향상 계층을 부호화하는 기술이다. HEVC 부호화기는 어떠한 블록에 대한 움직임 정보를 전송할 때, 부호화 비트를 줄이기 위하여 값을 전송하지 않고 후보로 선택된 주변 블록의 움직임 정보를 그대로 사용 (merge mode)하거나 주변 블록의 움직임 정보와의 차이를 전송 (Advanced Motion Vector Prediction, AMVP)한다[9]. 최적의 주변 블록을 결정하기 전에 <그림 9>와 같이 후보로 사용할 블록들을 탐색하며, 후보로 선택할 수 있는 총 블록의 개수는 5개이다. 현재 블록에 인접해있는 블록 중 4개를 spatial candidates로, 다른 시간에 위치한 프레임 내에 동일한 위치에 있는 블록 중 하나를 temporal candidate로 선택한다. <그림 9>에서 AL (Above Left), A (Above), AR (Above Right), L (Left), BL (Below Left) 블록은 현재 블록에 인접해있는 블록들이며 T는 동일한 계층에서 시간적으로 다른 프레임에 위치한 블록이다. SHVC에서는 하위 계층 내의 동일한 위치에 존재하는 블록 (C)의 움직임 정보 또한 temporal candidate로 사용될 수 있으며, T와 비교하여 예측 성능이 좋은 움직임 정보가 최종 temporal candidate로 쓰이게 된다. 공간적 스케일러빌리티에서는 계층 간 텍스처 예측과 동일하게, 계층 간 해상도의 차이를 고려하여 움직임 벡터 (Motion Vector, MV)를 업스케일 (Up-scale) 한 후에 사용한다.

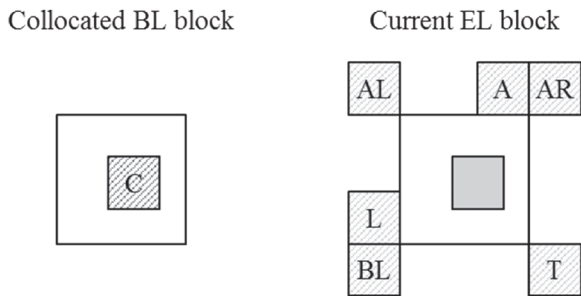


그림 9. SHVC에서의 Merge 및 AMVP 후보

## 2. SHVC Core Experiments (SCE)

2013년 7월에 열린 JCT-VC 14차 회의에서는 SCE를 3가지 카테고리로 분류하여 기술들의 성능을 검토하였다. 본 절에서는 각각의 SCE 카테고리에 대한 설명과 세부 기술 및 성능에 대해 알아보고, JCT-VC 15차 회의에서 진행 할 SCE 항목에 대해 간단히 살펴보도록 하겠다.

### 2.1 SCE 1: Support for additional resampling phase shifts

SCE 1은 공간적 스케일러빌리티에서 참조 계층을 업 샘플링 할 때, 다운 샘플러의 phase shifts를 고려한 업 샘플러를 테스트 하기 위한 SCE이다[10]. 현재 SHVC 표준화에서 실험을 위해 사용되고 있는 다운 샘플링 된 영상들은 <그림 10>과 같이 다운 샘플링 (zero-phase) 되며[11], SHVC에서 사용하고 있는 업 샘플링 필터 계수는 이와 같은 다운 샘플링 된 영상에 대해서만 고려되어 있다. 다운 샘플러는 SHVC 표준에 포함되지 않기 때문에, 만약 다운 샘플링이 <그림 10>과 같이 되어있지 않으면 참조 계층의 텍스처를 업 샘플링 해야 하는 공간적 스케일러빌리티에서 부호화 성능이 떨어질 수 있다. 이러한 문제점을 보완하기 위해서는 다운 샘플러의 phase shift를 고려한 업 샘플러를 사용해야 하며, SCE 1에서는 최적의 업 샘플러를 선정하기 위해 다양한 업 샘플링 필터에 대한 테스트를 진행하고 있다.

<그림 11>은 SCE 1과 관련된 기술들을 테스트하기 위해 사용한 다운 샘플러의 phase shift를 나타낸 것이다. 다운 샘플러의 phase를 고려한 업 샘플러의 성능을 살펴보면, Qualcomm에서 제안한 고정된 필터 계수를 사용하는 업 샘플러를 사용했을 때 SHM2.0 인코더 대비 BD-rate 성능이 약 6.7% 향상되었다[10].

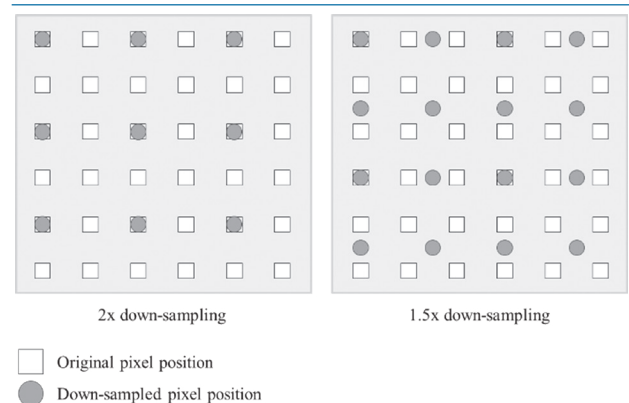


그림 10. 휘도 성분에 대한 다운 샘플링

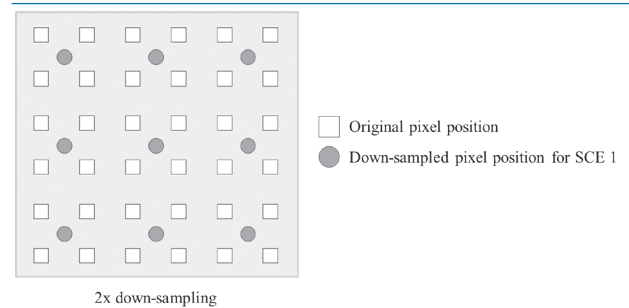


그림 11. SCE 1 테스트를 위한 다운 샘플링 방법

## 2.2 SCE 2: Combination of inter-layer syntax prediction and motion data compensation

앞서 계층 간 예측에서 설명한 바와 같이, 하위 계층의 움직임 정보와 같은 선택스 (syntax)를 참조하는 기술을 통틀어서 계층 간 선택스 예측 (inter-layer syntax prediction)이라고 한다. SCE 2에서는 하위 계층의 선택스 정보 중 하나인 움직임 정보를 사용하는 기술에 대해 테스트를 한다. HEVC 표준은 현재 프레임에 대한 코딩이 끝나면 프레임 내의 모든 움직임 정보를  $16 \times 16$  단위로 저장한다.  $8 \times 8$  또는  $8 \times 4$  등의 작은 블록 단위로 예측을 수행하였더라도 움직임 정보는  $16 \times 16$ 으로 압축 (compression)되어 저장된다. SHVC에서 향상 계층 부호화는 하위 계층 부호화가 끝난 후에 진행되기 때문에 향상 계층에서 참조하는 하위 계층의 모든 움직임 정보는  $16 \times 16$  단위로 압축된 형태이다. 따라서 계층 간 움직임 예측 시 움직임 정보의 정확도가 떨어질 수 있으며, SCE 2에서는 계층 간 움직임 예측과 하위 계층의 움직임 정보 저장 방법에 대한 성능을 분석해본다[12].

SCE 2에 대해서는 Canon, Sony 등의 회사에서 기술을 제안하였다. Canon은 하위 계층의 움직임 정보가 압축되는 정도에 따라 향상 계층 블록과의 대응되는 위치를 계산하는 방법을 제안하였고, 최대 0.15%의 BD-rate 성능 향상을 보였으며, SHVC 표준에 채택되었다. Sony는 하위 계층의 움직임 정보를 압축하는 것을 두 단계로 나누는 방법을 제안하였다. SHVC 부호화기는 시간적으로 동일한 시점에 있는 모든 계층의 프레임들을 부호화 한 후에 다음 시점에 대한 부호화를 진행한다. 즉, 하위 계층의 n번째 프레임을 코딩하고 향상 계층의 n번째 프레임을 코딩 한 후에 하위 계층의 n+1번째 프레임을 코딩 한다. 하위 계층의 n번째 프레임에 대한 움직임 정보는 향상 계층의 n번째 프레임을 모두 부호화 한 후에 압축하여도 되기 때문에 향상 계층에서는 압축되지 않은 하위 계층의 움직임 정보를 사용할 수 있다. 이러한 점을 이용하여, Sony에서는 하위 계층 코딩이 끝나면 움직임 정보를  $16 \times 16$ 으로 압축하지 않고  $8 \times 8$  단위로 압축을 하여 이를 향상 계층에서 참조하는 구조를 제안하였다. 향상 계층 프레임 코딩이 끝나면 하위 계층의 움직임 정보는  $16 \times 16$ 으로 압축된다. 이 기술은 BD-rate 성능을 최대 0.44% 향상시킬 수 있지만 추가적인 메모리가 요구된다는 단점이 있다.

## 2.3 SCE 3: Inter-layer filtering

SCE 3는 계층 간 필터링 (inter-layer filtering) 관련 기술로, 참조 계층 픽처를 향상시키기 위해 취해지는 필터들을 테스트 한다[13]. 필터는 참조 계층을 향상 계층의 참조 리스트에 넣기 전에 적용되며, 공간적 스케일러빌리티의 경우 참조 계층

을 향상 계층 해상도로 업 샘플링 한 후에 계층 간 필터링을 적용한다. 계층 간 필터링의 종류에는 픽처 또는 블록 단위의 적응적 루프 필터 (adaptive loop-filter), 샤프닝 필터 (sharpening filter), 휘도 샘플을 이용한 색차 성분 필터 (chroma filter) 등이 있다. 휘도 성분에 대한 필터의 성능을 보면 대부분은 화질적 스케일러빌리티에서 성능이 높게 나타난다. 공간적 스케일러빌리티에서는 0.1 ~ 0.5% 정도의 BD-rate 성능 향상이 있지만, 화질적 스케일러빌리티에서는 1.5 ~ 2.0% 정도 부호화 성능이 향상된다. 삼성에서 제안한 색차 성분 필터의 경우 공간적 스케일러빌리티에서는 8%, 화질적 스케일러빌리티에서는 6% 정도의 BD-rate 성능 향상을 보였다. JCT-VC 14 차 회의에서 SCE 3과 관련하여 채택된 기술은 없으며, 다음 회의에서 inter-layer filtering에 대한 SCE를 계속 진행할 것으로 보인다.

## 2.4 JCT-VC 15차 회의를 위한 SCE

JCT-VC 14차 회의가 마무리되면서 다음 회의에서 논의할 SCE 주제를 결정하였다. SCE는 14차 회의와 마찬가지로 총 3개의 카테고리 나뉘어졌으며, SCE 1과 SCE 2는 새로운 주제로 바뀌었고 SCE 3는 계속 진행하는 것으로 결정되었다. SCE 항목 및 요약 내용은 <표 3>과 같으며, 각각의 SCE에 대한 세부적인 기술은 JCT-VC 문서를 통해 확인 할 수 있다[14][15][16].

표 3. SCE 요약

No.	Category	Summary
1	Arbitrary scalability ratio support	2배, 15배 외에 다양한 공간적 스케일러빌리티에 대한 다운 샘플링 및 업 샘플링
2	Key pictures and single-loop decoding	단일-루프 (single-loop) 구조 및 key picture concept에 대한 scheme
3	Inter-layer filtering	계층 간 텍스처 예측 성능 향상을 위한 필터링 기술

## VI. SHVC 성능 평가

본 장에서는 SHVC의 부호화 성능을 알아본다. 실험에는 두 개의 계층 (참조 계층, 향상 계층)을 사용하였으며, HEVC기반 SHVC 소프트웨어인 SHM2.0으로 부호화 성능 실험을 진행하였다. 부호화 성능 실험은 SHVC 공통 실험 조건[6] 하에 진행되었으며, 실험에 사용된 영상 및 기본 계층과 향상 계층에 대한 QP 값 설정은 표 4와 같다.



〈표 5〉는 향상 계층 영상에 대해 simulcast 방식처럼 HEVC로 독립적으로 부호화한 경우와 계층 간 예측 툴을 사용하여 부호화하는 SHVC 경우에 대한 성능을 BD-rate 값으로 비교한 것이다[6]. 이때 기본 계층은 두 경우 모두 HEVC로 독립적으로 부호화하는 것과 동일 하기 때문에 향상 계층의 부호화 성능만 비교하였다. SHVC의 공간적 스케일러블 부호화 성능을 살펴보면, 2배 공간적 스케일러블 부호화 성능은 약 27% 이고, 1.5배의 부호화 성능은 약 48% 이다. 1.5배가 2배 대비 21% 정도 부호화 성능 향상이 높은 것을 알 수 있다. 일반적으로 2배 다운샘플링 된 영상에 비해 1.5배 다운샘플링 된 영상의 텍스처가 향상 계층과의 유사성이 높기 때문에 계층 간 예측 툴에 의한 부호화 성능 향상이 더 높게 나타난다. 화질적 스케일러블 부호화의 경우, simulcast로 부호화 한 것 대비 26%의 부호화 성능 향상이 있다. 화질적 스케일러빌리티는 모든 계층 영상의 해상도가 동일하지만 부호화 성능은 1.5배 공간적 스케일러블 부호화 대비 21% 정도 낮다. 이는 화질적 스케일러빌리티의 경우 계층 간 화질 (QP)의 차이가 있기 때문에 동일한 해상도임에도 불구하고 계층 간 텍스처의 유사성이 크게 떨어지기 때문이라고 볼 수 있다.

표 4. SHVC 실험 영상 및 QP

Scalability	Seq.	Resolution		QP	
		BL	EL	QP <sub>BL</sub>	QP <sub>EL</sub>
2x spatial scalability	Class A	1280×800	2560×1600	22 26 30 34	QP <sub>BL</sub> -0 QP <sub>BL</sub> +2
	Class B	960×540	1920×1080		
1.5x spatial scalability	Class A	-	-		
	Class B	1280×720	1920×1080		
SNR scalability	Class A	2560×1600		26 30 34 38	QP <sub>BL</sub> -6 QP <sub>BL</sub> -4
	Class B	1920×1080			

표 5. HEVC 대비 SHVC 향상 계층 부호화 성능

(단위:%)	AI		RA			LD-B		
	2x	1.5x	2x	1.5x	SNR	2x	1.5x	SNR
Class A	-42.2	-	-29.3	-	-33.6	-19.5	-	-22.6
Class B	-31.5	-57.3	-23.9	-46.7	-29.9	-15.5	-39.3	-18.3
Avg.	-36.8	-57.3	-26.6	-46.7	-31.8	-17.5	-39.3	-20.4

## Ⅷ. 결론

본 고에서는 스케일러빌리티의 개념 및 HEVC를 기반으로 하는 스케일러블 영상 압축 표준인 SHVC에 대해 알아보았다. JCT-VC는 현재 SHVC 부호화 기술들을 평가하기 위해 SCE를 진행하고 있으며, SCE 관련 기술 이외에도 SHVC의 압축 성능 향상을 위한 기술들과 메모리 감소와 알고리즘 간소화와 같은 기술들도 기고되고 있다. 이러한 SHVC 표준은 2014년 6월에 표준화가 완료될 예정이며, SHVC는 향후 멀티미디어 시장이 확대되고 다양한 서비스 환경이 구축됨에 따라 콘텐츠 서비스를 위한 중요한 기술 중 하나로 자리매김 할 것이라 예상된다.

## 참고 문헌

- [1] Heiko Schwarz and Detlev Marpe, and Thomas Wiegand “Overview of the scalable video coding extension of the H.264/AVC standard,” IEEE Transactions on Circuits and Systems for Video Technology, 2007.10
- [2] ISO/IEC JTC 1/SC 29/WG 11 and ITU-T SG 16 WP 3 “Joint Call for Proposals on Scalable Video Coding Extensions of High Efficiency Video Coding (HEVC),” ISO/IEC JTC1/SC29/WG11 N12957, 2012.7
- [3] Jung Won Kang, Hahyun Lee, Jinho Lee, Jin Soo Choi, Jin Woong Kim, Junghak Nam, Hyomin Choi and Donggyu Sim “Description of scalable video coding technology proposal by ETRI and Kwangwoon Univ.,” Document of JCT-VC (JCTVC-K0037), 2012.10
- [4] Hyomin Choi, Junghak Nam and Donggyu Sim “Scalable structures and inter-layer predictions for HEVC scalable extension,” Document of JCT-VC (JCTVC-F096), 2011.7
- [5] Jianle Chen, Jill Boyce, Yan Ye, and Miska M. Hannuksela “SHVC Test Model 2 (SHM2),” Document of JCT-VC (JCTVC-M1007), 2013.4
- [6] Xiang Li, Jill Boyce, Patrice Onno, and Yen Ye “Common SHM test conditions and software reference configurations,” Document of JCT-VC (JCTVC-M1009), 2013.4



- [7] Hyomin Choi, Junghak Nam, Donggyu Sim, and Ivan V. Bajic "Scalable video coding based on high efficiency video coding (HEVC)," IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, 2011.8
- [8] Jianle Chen, Jill Boyce, Yan Ye, and Miska M. Hannuksela "SHVC Working Draft 2," Document of JCT-VC (JCTVC-M1008), 2013.4
- [9] Benjamin Bross, Woo-Jin Han, Jens-Rainer Ohm, Gary J. Sullivan, Ye-Kui Wang, and Thomas Wiegand "High Efficiency Video Coding (HEVC) text specification draft 10 (for FDIS & Last Call)," Document of JCT-VC (JCTVC-L1003), 2013.1
- [10] Elena Alshina, Xiang Li, and Jie Dong "SCE1: Summary Report of SHVC Core Experiment on support for additional re-sampling phase shifts," Document of JCT-VC (JCTVC-N0031), 2013.8
- [11] InterDigital Communications, LLC "Downsampling filters for anchor generation for scalable extensions of HEVC," ISO/IEC JTC1/SC29/WG11 MPEG2012/M24499, 2012.5
- [12] Christophe Gisquet and Kazushi Sato "SCE2: Summary Report of SHVC Core Experiment on combination of inter-layer syntax prediction and motion data compression," Document of JCT-VC (JCTVC-N0032), 2013.8
- [13] Jianle Chen, Andrew Segall, Elena Alshina, Shan Liu, and Jie Dong "SCE3: Summary Report of SHVC Core Experiment on Inter-layer Filtering," Document of JCT-VC (JCTVC-N0033), 2013.8
- [14] Edouard Francois, Elena Alshina, and Jianle Chen "Description of HEVC Scalable Extension Core Experiment SCE1: Arbitrary scalability ratio support," Document of JCT-VC (JCTVC-N1101), 2013.8
- [15] Mathias Wien, Krishna Rapaka, and Xiaoyu Xiu "HEVC Scalable Extension Core Experiment SCE2: Key pictures and single-loop decoding," Document of JCT-VC (JCTVC-N1102), 2013.8
- [16] Jianle Chen, Elena Alshina, Jie Dong, and Maxim Sychev "Description of HEVC Scalable Extension Core Experiment SCE3: Inter-layer filtering,"

Document of JCT-VC (JCTVC-N1103), 2013.8

## 약 력



심 동 규

1999년 서강대학교 전자공학과 박사  
1999년~2000년 (주) 현대 전자  
2000년~2002년 (주) 바로 비전  
2002년~2005년 Univ. of Washington  
2005년~현재 광운대학교 컴퓨터공학과 교수  
관심분야: 영상처리, 영상압축, 컴퓨터비전



김 경 혜

2012년 광운대학교 컴퓨터공학과 학사  
2012년~현재 광운대학교 컴퓨터공학과 석사과정  
관심분야: 영상처리, 스케일러블 비디오 코딩



조 현 호

2008년 광운대학교 컴퓨터공학과 학사  
2010년 광운대학교 컴퓨터공학과 석사  
2010년~현재 광운대학교 컴퓨터공학과 박사과정  
관심분야: 영상처리, 영상압축